DATASET
Dataset is sample data of songs heard by users on an online streaming platform. The Description of data set attached in musicdata.txt is as follows: -
1st Column - UserId
2nd Column - TrackId
3rd Column - Songs Share status (1 for shared, 0 for not shared)
4th Column - Listening Platform (Radio or Web - 0 for radio, 1 for web)
5th Column - Song Listening Status (0 for skipped, 1 for fully heard)

111115|222|0|1|0
111113|225|1|0|0
111117|223|0|1|1
111115|225|1|0|0

Task 2 : What are the number of times a song was heard fully.

Here we need to find the number of times which track played how many times.

NOTE:
The above data set is already present under haddop fs in the path "**/satya/MR/song.txt**"

```
// Assignment5Task2.java

import java.io.IOException;
import java.util.ArrayList;

import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.Reducer.Context;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class Assignment5Task2 {


        public static class Assignment5Task2Mapper extends Mapper<Object, Text, Text, IntWritable>{

                private final static IntWritable mapValue = new IntWritable(1);
                private Text mapKey = new Text();
```

```
		public void map(Object key, Text value, Context context) throws IOException,
InterruptedException {
			//StringTokenizer itr = new StringTokenizer(value.toString());
			//while (itr.hasMoreTokens()) {
			//word.set(itr.nextToken());
			//context.write(word, one);
			// }
			//}
			System.out.println("START#map()");
			String [] values= value.toString().split("\\|");

			if(values!=null && values.length==5) {
				String trackId = values[1];
				System.out.println("Track ID ::"+trackId);


				String listeningStatus = values[4];
				System.out.println("Listening Status ::"+listeningStatus);

				if(trackId!=null && trackId.trim().length()>0
						&& listeningStatus!=null &&
listeningStatus.trim().length()>0
							&& Integer.parseInt(listeningStatus)>0)
				{
					System.out.println("As the condition match, so adding this to
Context Object");
					mapKey.set(trackId);
					context.write(mapKey,mapValue);
				}

			}
			System.out.println("END#map()");
		}
	}

	public static class Assignment5Task2Reducer extends
Reducer<Text,IntWritable,Text,IntWritable> {
		private IntWritable result = new IntWritable();

		public void reduce(Text key, Iterable<IntWritable> values,
				Context context
				) throws IOException, InterruptedException {

			System.out.println("START#reduce()");

			int noOfTimesListen = 0;

			for (IntWritable val : values) {
```

```
                        noOfTimesListen += val.get();
                    }
                    result.set(noOfTimesListen);
                    context.write(key, result);
                    System.out.println(" KEY "+key.toString() + " : Value "+noOfTimesListen);
                    System.out.println("END#reduce()");
                }
        }

         @SuppressWarnings("deprecation")
        public static void main(String[] args) throws Exception {
                 //create an instance of Configuration object
           Configuration conf = new Configuration();
           conf.addResource(new Path("/home/acadgild/install/hadoop/hadoop-
2.6.5/etc/hadoop/core-site.xml"));
                conf.addResource(new Path("/home/acadgild/install/hadoop/hadoop-
2.6.5/etc/hadoop/hdfs-site.xml"));

                //creatkeye an instance of FileSystem that holds Filesystem namespace
                 FileSystem fs = FileSystem.get(conf);



            System.out.println("Usage: song <input file> <output dir>");
            System.out.println("Using default file: song.txt");
          //variables to hold path of input file and output directory

           // HDFC FILE PATH
            String inPath = "/satya/MR/song.txt";
            String outPath = "/satya/MR/Output/Task2";
            //Normal File System
            //String inPath = "/home/acadgild/Desktop/MyDocument/read/wordcount.txt";
            //String outPath = "/home/acadgild/Desktop/MyDocument/read/WordCountOutput2";

        //create an instance of job
                try {
           Job job = new Job(conf, "Music Count Task-2");
           job.setJarByClass(Assignment5Task2.class);
           job.setMapperClass(Assignment5Task2Mapper.class);
           job.setReducerClass(Assignment5Task2Reducer.class);

           job.setNumReduceTasks(1);

           job.setMapOutputKeyClass(Text.class);
           job.setMapOutputValueClass(IntWritable.class);

           job.setOutputKeyClass(Text.class);
           job.setOutputValueClass(IntWritable.class);
```

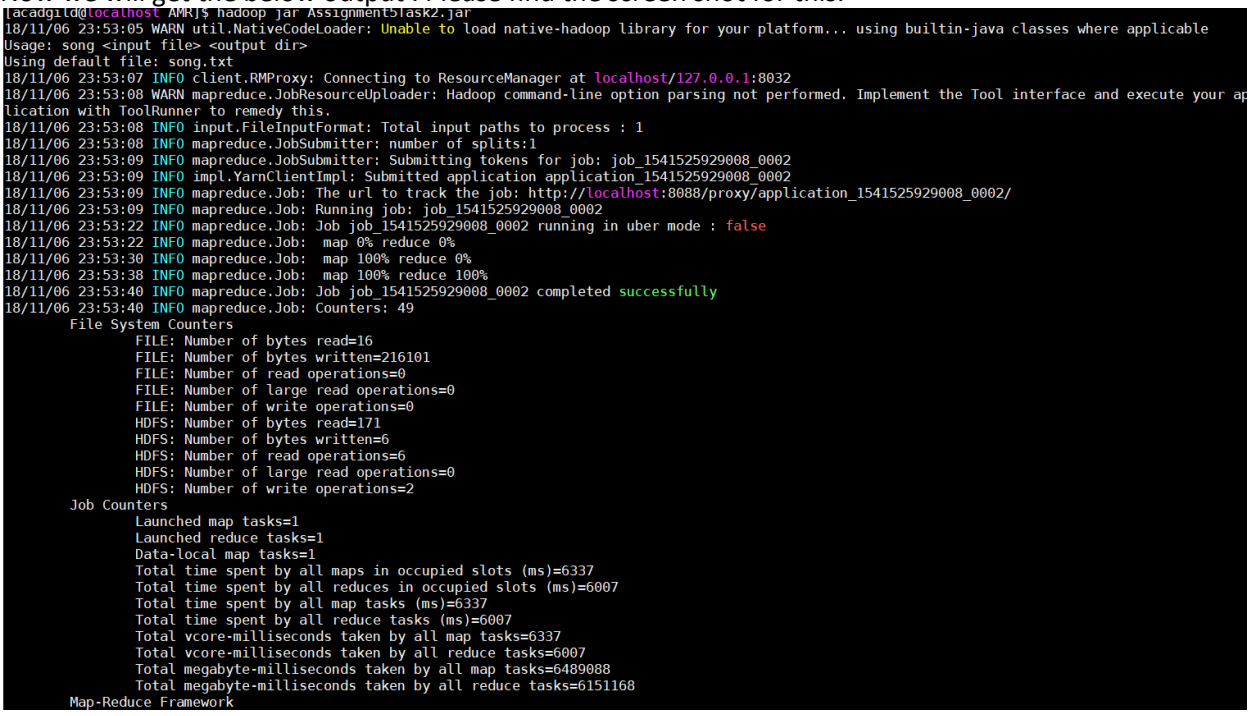
```
        FileInputFormat.addInputPath(job, new Path(inPath));

        if (fs.exists(new Path(outPath))) {
                fs.delete(new Path(outPath), true);
        }
        FileOutputFormat.setOutputPath(job, new Path(outPath));
        System.exit(job.waitForCompletion(true) ? 0 : 1);
        }catch(Exception e) {
                System.out.println(e);
        }
    }
```

}

Now extract the jar file in to the local folder **/home/acadgild/Desktop/Practise/AMR/ Assignment5Task2.jar** and run it with the help of below command.

hadoop jar Assignment5Task2.jar

Now we will get the below output . Please find the screen shot for this.



```
[acadgild@localhost AMRJ]$ hadoop jar Assignment5Task2.jar
18/11/06 23:53:05 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Usage: song <input file> <output dir>
Using default file: song.txt
18/11/06 23:53:07 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/11/06 23:53:08 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your ap
lication with ToolRunner to remedy this.
18/11/06 23:53:08 INFO input.FileInputFormat: Total input paths to process : 1
18/11/06 23:53:08 INFO mapreduce.JobSubmitter: number of splits:1
18/11/06 23:53:09 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1541525929008_0002
18/11/06 23:53:09 INFO impl.YarnClientImpl: Submitted application application_1541525929008_0002
18/11/06 23:53:09 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1541525929008_0002/
18/11/06 23:53:09 INFO mapreduce.Job: Running job: job_1541525929008_0002
18/11/06 23:53:22 INFO mapreduce.Job: Job job_1541525929008_0002 running in uber mode : false
18/11/06 23:53:22 INFO mapreduce.Job:  map 0% reduce 0%
18/11/06 23:53:30 INFO mapreduce.Job:  map 100% reduce 0%
18/11/06 23:53:38 INFO mapreduce.Job:  map 100% reduce 100%
18/11/06 23:53:40 INFO mapreduce.Job: Job job_1541525929008_0002 completed successfully
18/11/06 23:53:40 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=16
                FILE: Number of bytes written=216101
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=171
                HDFS: Number of bytes written=6
                HDFS: Number of read operations=6
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=6337
                Total time spent by all reduces in occupied slots (ms)=6007
                Total time spent by all map tasks (ms)=6337
                Total time spent by all reduce tasks (ms)=6007
                Total vcore-milliseconds taken by all map tasks=6337
                Total vcore-milliseconds taken by all reduce tasks=6007
                Total megabyte-milliseconds taken by all map tasks=6489088
                Total megabyte-milliseconds taken by all reduce tasks=6151168
        Map-Reduce Framework
```

```
        Map-Reduce Framework
                Map input records=4
                Map output records=1
                Map output bytes=8
                Map output materialized bytes=16
                Input split bytes=104
                Combine input records=0
                Combine output records=0
                Reduce input groups=1
                Reduce shuffle bytes=16
                Reduce input records=1
                Reduce output records=1
                Spilled Records=2
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=135
                CPU time spent (ms)=1540
                Physical memory (bytes) snapshot=286388224
                Virtual memory (bytes) snapshot=4118216704
                Total committed heap usage (bytes)=170004480
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=67
        File Output Format Counters
                Bytes Written=6
u have new mail in /var/spool/mail/acadgild
cadgild@localhost AMR]$
```

Now we will run the cat command to see the output.
As my output directory is **"/satya/MR/Output/Task2", we can run below command to see the output**
hadoop fs -cat /satya/MR/Output/Task2/p*

```
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost AMR]$ hadoop fs -cat /satya/MR/Output/Task2/p*
18/11/06 23:55:38 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
223     1
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost AMR]$
```